

# REGION-WISE PAGE DIFFICULTY ANALYSIS USING EYE MOVEMENTS

Tsubasa Minematsu  
*Kyushu University*  
744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan

## ABSTRACT

In this study, we investigated which section of a page was difficult for students to read, based on eye movement data and subjective impressions of the page's difficulty, with the aim of helping teachers revise teaching materials. It is problematic to manually model relationships between eye movements and subjective impressions of the page's difficulty. Therefore, in this study, we used a neural network to model the relationships automatically. Our method generated relevance maps representing locations where students found difficulty, in order to visualize region-wise page difficulty. To evaluate the quality of the relevance maps, we compared them with a distribution of gaze points and highlights added by the students. In addition, we administered a questionnaire to evaluate whether the relevance maps were useful to teachers when revising teaching materials. Results imply that our method can provide useful information for teachers making revisions to teaching materials.

## KEYWORDS

Learning Analytics, Eye Movement, Machine Learning, Neural Network

## 1. INTRODUCTION

Improvement of learning supports is vital for enhancing education. Knowing what students find difficult is useful for improving learning supports. If teachers know what students are not able to understand in their lectures, it is easier for them to revise their teaching materials and thus teach difficult contents more carefully. In addition, this information can help to improve the summarization of current teaching materials (Shimada, et al., 2015) and inform recommendations for supplemental materials (Shiino, et al., 2019) provided by e-learning systems.

Several researchers have focused on relationships between learning behaviors and students' subjective impressions of difficulty. Nakamura et al. predicted subjective impressions of difficulty of English word tests by combining features such as eye movements and head poses (Nakamura, et al., 2008). Ohkawauchi et al. (Ohkawauchi, et al., 2012) and Shiino et al. (Shiino, et al., 2019) focused on more complex teaching materials, such as textbooks containing figures, tables, text, formulations, and images. Ohkawauchi et al. asked students to rate the degrees of difficulty in specific sections of teaching materials. The difficulty ratings were shown to teachers as evidence of students' understanding. Shiino et al. estimated the difficulty of each page of certain teaching materials by using students' clickstream data recorded from an e-learning system M2B (Ogata, et al., 2015).

These previous studies did not consider where on the page students found difficulty. The location information is more useful than teaching material-wise information and page-wise information. However, it is difficult to collect the location information by using questionnaires. To solve this problem, eye movement data can be used. The analysis of eye movement data can be helpful in understanding learning behaviors within pages. In previous works using eye movement data, findings have been shown related to effective attention guidance techniques (De Koning, et al., 2010); effectiveness of using both text and pictures in teaching materials (Mason, et al., 2013); and relationships between students' scan paths and performance (The & Mavrikis, 2016) (Jian & Ko, 2017). We believe that eye movement data can be helpful in representing how students comprehend the contents of different teaching materials.

In this study, we investigated which section of a page was difficult for students to comprehend based on the students' eye movements while studying by themselves. The research reported in this paper attempted to model relationships between subjective impressions of difficulty and eye movement data. We then visualized page regions related to the difficulty reported. We used a neural network for modeling such relationships because it is difficult to manually design features for representing eye movements. Visualization of a difficult page was performed based on finding the network's neurons related to the difficulty. In our experiment, we evaluated sections of a page students found difficult, with the aim of supporting teachers when they revise their materials in the future.

Our research questions are summarized as follows:

R1. Can we model relationships between subjective impressions of difficulty and eye movement data?

R2. Is our visualization of the page regions where students find difficulty useful for teachers when they revise textbooks?

## 2. DATA COLLECTION

We performed all procedures in accordance with the approved guidelines of the ethics committee of Kyushu University. In addition, we received prior written informed consent from participants in accordance with the Declaration of Helsinki.

We focused on analyzing reading patterns within pages of teaching materials from the eye movements of students. The eye movement data were collected from 15 university students engaged in an e-learning system. The 15 participants were undergraduate students in Kyushu University (seven females) with a mean age of 20.2 years ( $SD=1.6$ ). Before our experiment, we confirmed that all participants had little knowledge or experience regarding information science and the statistical mathematics presented in our experiment. In this study, we used a Tobii eye tracker (Tobii pro spectrum 150 Hz) which was attached to a monitor. The monitor displayed teaching materials; the distance between the monitor and the eyes of the students was 57 cm. We measured their eye movements in a dark room with each student individually in order to reduce the effects of ambient noise. In our data collection, the sampling rate was set to 150 Hz.

Before the measurement, we calibrated the eye tracker device for each student. After completing the calibration, students viewed teaching materials in the e-learning system, consisting of a statistical test and correlation. The contents included figures, tables, text, formulations, and images. In addition, the content alignment was free.

We asked all the students to give their subjective impressions of a page's difficulty. Students used a slider interface to provide an impression score on a scale of zero to ten after finishing each page. The higher the score, the more difficult the page. To initialize eye movements, a black page was displayed for one second before the students started to read the next page. In addition, students could read previous pages freely.

Students were made to participate in an examination after they finished reading all the pages. In order to enhance their motivation, we informed the students about the examination before the measurement. In addition, the students added highlights on each page where they found difficult contents. Finally, our measurement covered the students' eye movements, subjective impressions of page difficulty, and highlights.

### 2.1 Preliminary Analysis

We show distributions of subjective impressions of page difficulty for each student. As shown in Figure 1, subjective impressions of the page difficulty were distributed, and the means and variances were different between the students. As we know, it is difficult to analyze such subjective impressions based on absolute values. In this study, we normalized subjective impressions of page difficulty for each student between zero and one. The normalized values were binarized using a threshold value. If the normalized value of the student was more than the threshold value, it was classified as a difficult page. We set the threshold value to 0.75 experimentally.

Students' subject impressions of page difficulty

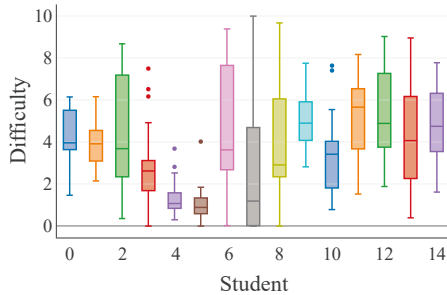


Figure 1. Distribution of subjective impression of page difficulty

Histogram of the number of students finding difficulty

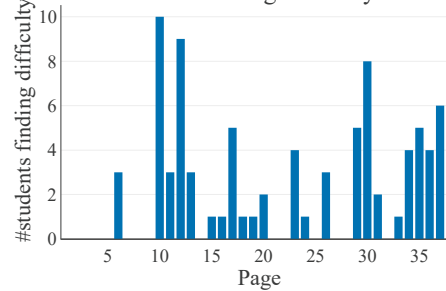


Figure 2. Distribution of the number of students finding difficulty in each page

Figure 2 shows the distribution of the number of students finding difficulty in each page. We could confirm that more than half of the students felt some specific pages to be difficult, such as pages 10, 12, and 30. These pages contained explanations about the definition of correlation coefficient and an example of a t-test. Based on this preliminary analysis, we observed that some students found the contents of the teaching materials difficult.

### 3. READING PATTERN ANALYSIS BASED ON A NEURAL NETWORK

#### 3.1 Modeling Relationship between Students' Eye Movement and Difficulty

We used a neural network to investigate the relationship between students' eye movements and subjective impressions of a page's difficulty. The neural network used in this study accepted eye movement data and then classified whether the data were related to difficult pages. The advantage of neural networks is that they obtain effective features for classification automatically based on input data. In other machine learning approaches, we have had to design effective features manually. However, it was difficult to design manually in our study because we focused on page difficulty and student's eye movements, which represented both temporal and spatial information. Therefore, we chose to use a neural network.

Our neural network accepted a three-dimensional tensor as the input and provided a probability of page difficulty. The probability of page difficulty represented whether a tensor generated from eye movement data belonged to a difficult page. The tensor represented a student's eye movement and was generated based on reading pattern codes proposed by Minematsu et al. (Minematsu, et al., 2019). We encoded a sequence of eye movement data to three-dimensional tensors, which represented temporal and spatial information. First, the sequence was divided into  $T$  time slots. Then, we computed a density map of gaze points at each time slot based on a kernel density estimation. The size of the density map was  $H \times W$ . Therefore, we obtained a  $H \times W \times T$  tensor from the sequence of eye movements on the page. The first and second axes represent spatial information, and the third axis represents temporal information. We did not form a one-dimensional vector because convolution layers can accept high dimensional tensors.  $H$ ,  $W$ , and  $T$  were set to 20 experimentally.

Our neural network architecture is described in Table 1. Our neural network contained five convolution layers, two fully connected layers, and three max-pooling layers. In all convolution layers and fully connected layers, the kernel size was  $3 \times 3$ , and the stride was  $1 \times 1$ . In all max-pooling layers, the kernel size was  $2 \times 2$ , and the stride was  $2 \times 2$ . We used rectified linear units (ReLUs) as activation functions, with the exception of the output layer. In the output layer, a sigmoid function was used to provide a probability of difficult pages. We used cross entropy cost function to train our neural network. Our neural network was optimized using the Adam optimizer based on backpropagation. The learning rate was set to 0.0001. The other parameters were set to the default values described in (Kingma & Ba, 2015). After training, our neural network was able to model relationships between eye movements and difficult pages.

Table 1. Neural network architecture

Layer	Input shape (height × width × channel)	Output shape (height × width × channel)
Input	20 × 20 × 20	20 × 20 × 20
Convolution	20 × 20 × 20	20 × 20 × 32
Convolution	20 × 20 × 32	20 × 20 × 32
Max-pooling	20 × 20 × 32	10 × 10 × 32
Convolution	10 × 10 × 32	10 × 10 × 64
Max-pooling	10 × 10 × 64	5 × 5 × 64
Convolution	5 × 5 × 64	5 × 5 × 64
Max-pooling	5 × 5 × 64	3 × 3 × 64
Convolution	3 × 3 × 64	1 × 1 × 64
Fully connected	1 × 1 × 64	1 × 1 × 64
Fully connected	1 × 1 × 64	1 × 1 × 1

### 3.2 Interpretation of the Model

We used layer-wise relevance propagation (LRP) (Lapuschkin, et al., 2016) to visualize page areas where students found difficulty. LRP can provide a relevance score in each element of an input tensor of eye movements. The relevance score represents the contribution to the decision by our neural network at each element of the input tensor. When a relevance score in an element is positive, it means that the element supports the decision. Therefore, we focused on positive relevance scores in each input tensor to analyze which part of eye movement was related to the subjective impression of page difficulty. The details are referred to in (Lapuschkin, et al., 2016).

To summarize difficult areas on each page, we integrated the relevance scores of all students. First, relevance scores in each input tensor were normalized between -1 and 1. The normalization was performed by dividing each relevance score by the maximum value of the absolute value. Second, we summed the relevance tensors along the third axis, which represented temporal information, and then summed the relevance maps of all students. After the summation, we obtained relevance maps for each page. The size of the relevance maps were  $H \times W$ . The relevance maps were normalized again between -1 and 1. Finally, we extracted positive values, and applied the following function to extract values, in order to clarify the magnitude of relevance scores. The function is follows:

$$f(x) = \frac{1.0}{(1 + \exp(-10(x - 0.5)))} \quad (1)$$

where  $x$  is a relevance score. In the relevance map, a region with a large score contributes to classifying the page as a difficult page. In other words, the region can be related to difficult contents within the page.

## 4. EXPERIMENT

We obtained relevance maps by applying the method described in Section 3. The relevance maps were used to understand where students found difficulty on the page. We also visualized a distribution of gaze points for each page, in order to understand where students were looking. This distribution was called a gaze map in this study. The gaze map of the page was generated from all of the input tensors of eye movements on the page by following the same procedure as that used to generate relevance maps. In the gaze maps, when many students looked at a region for a long time, the region had a large value. In addition, the highlights added to the pages by students were available. We used the highlight maps to represent location and number. To obtain the highlight maps, we counted the number of highlights added on the same location of each page, which was then divided by the maximum value in all pages. We evaluated the relevance maps using the gaze maps and the highlight maps.

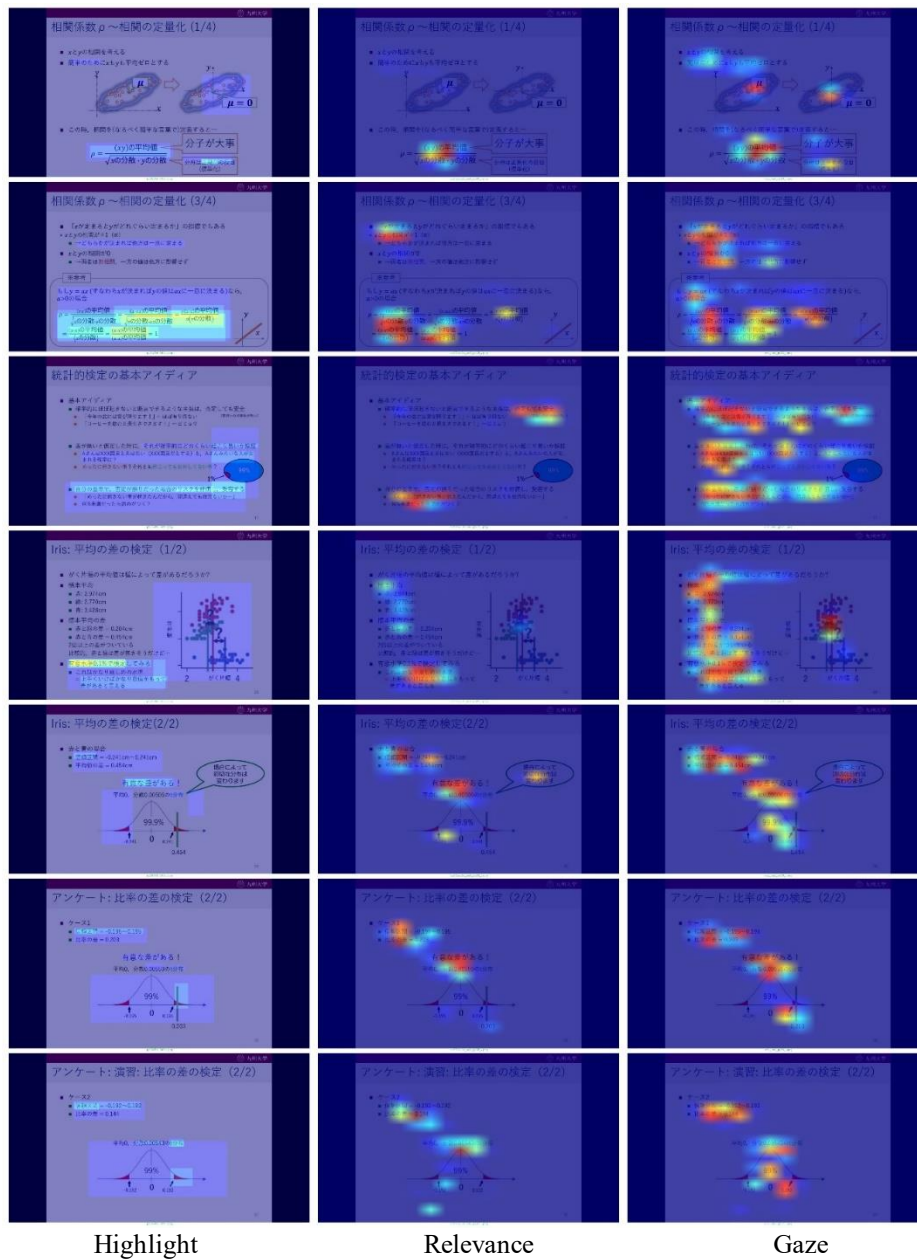


Figure 3. Visualization of highlight maps, relevance maps, and gaze maps in pages with more than five students who found difficulty on those pages

## 4.1 Qualitative Evaluation

We compared the relevance maps with the gaze maps and the highlight maps. Figures 3 and 4 show the highlight maps, the relevance maps, and the gaze maps superimposed on the corresponding pages. Figure 3 only includes those with more than five students who found difficulty on those pages, while Figure 4 shows the remainder. A figure in red has a larger value than a figure in blue.

According to the highlight maps in Figures 3 and 4, we could roughly confirm where some students found difficulty. The relevance maps in Figure 3 support this conclusion more than those in Figure 4. For example, the relevance maps and the highlight maps focus on equations in the first and the second row of Figure 3. We believe that it is easy for the neural network to model the eye movements and subjective impressions of the pages with contents the students found difficult.

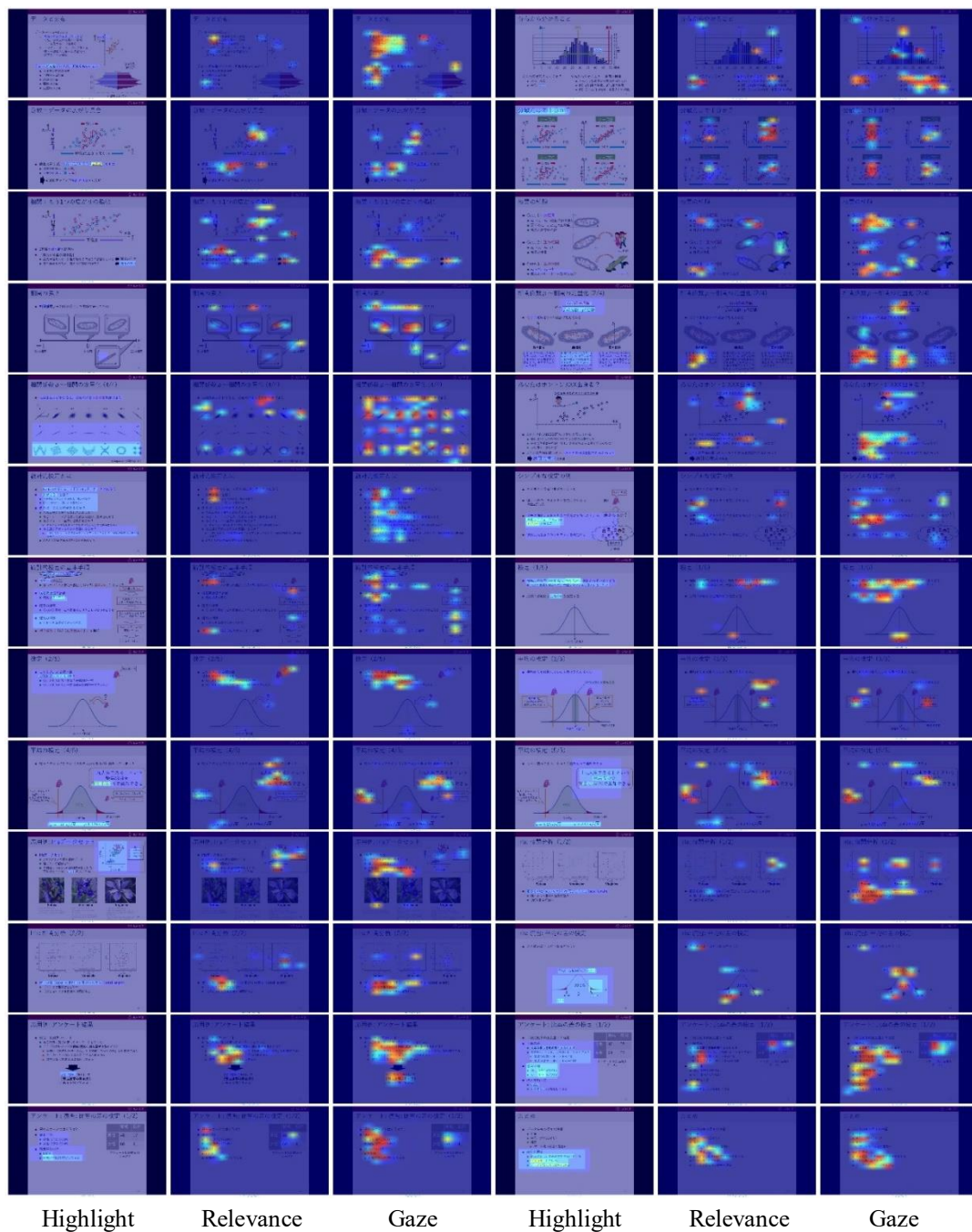


Figure 4. Visualization of highlight maps, relevance maps, and gaze maps in pages with less than four students who found difficulty on those pages

The gaze maps showed where students looked most frequently. Comparing the relevance maps with the gaze maps, the relevance maps showed more specific regions than the gaze maps. The neural network accepted tensors of eye movements as the input. Some relevance scores in the tensors became negative when performing LRP. Therefore, scores in the relevance maps were more limited than scores in the gaze maps.

## 4.2 Evaluation for Modification of Teaching Materials

We administered a questionnaire to the creator of the teaching materials in order to evaluate the quality of the relevance map from the point of view of a teacher. In Section 4.1, we confirmed that the relevance maps could be similar to the results in the highlight maps. However, the relevance maps may not always help teachers revise teaching materials. To investigate whether the relevance maps support teachers, we showed the relevance maps and the gaze maps to the creator of the teaching materials. Note that we did not explain how the maps were generated; only that they were generated by two different systems. In fact, system A generated the gaze maps, and system B generated the relevance maps.

Table 2 shows the questionnaire about the relevance maps and the gaze maps, and the responses. Q1 asked what the creator found to be difficult for the students. Q2 asked whether the creator would refer to the maps when modifying teaching materials. The creator answered the questions for each page.

According to the answer to Q1, the creator believed that the gaze maps represented page difficulty more accurately than the relevance maps. However, the answer to Q2 showed that the creator did not choose both the gaze maps and the relevance maps for every page. We believed the result was related to the page difficulty. To confirm, we focused on more difficult pages. If the number of students finding difficulty in a page is less than the threshold values, we ignore the page when computing the weighted average. Table 3 shows some weighted averages when easier pages were removed. In Q2, we confirmed that almost all weighted averages increased. This meant that the creator tended to refer to the relevance maps in difficult pages. Therefore, the relevance maps were more useful for the creator than the gaze maps in the modification of teaching materials.

Table 2. Questionnaire about the relevance maps and the gaze maps

Question		Evaluation					
		Agree A (1)	Agree A a little (2)	Neither agree nor disagree (3)	Agree B a little (4)	Agree B (5)	Weighted Average (n =33)
Q1	Which of the systems present similar results to what you find difficult for the students?	11	6	8	6	2	2.45
Q2	Which system do you want to refer to when modifying teaching materials?	1	4	22	3	3	3.09

Table 3. Weighted average varying threshold values about the number of difficult pages. n is the number of pages used for computing their weighted average

Question	Threshold values							
	$\geq 0$ (n=33)	$\geq 1$ (n=23)	$\geq 2$ (n=17)	$\geq 3$ (n=15)	$\geq 4$ (n=11)	$\geq 5$ (n=7)	$\geq 6$ (n=4)	$\geq 7$ (n=3)
Q1	2.45	2.26	2.35	2.47	2.55	2.57	2.00	2.00
Q2	3.09	3.13	3.18	3.27	3.36	<b>3.57</b>	3.25	3.33

## 5. DISCUSSION

In this study, we visualized the locations of difficult contents on the relevance maps by modeling relationships between students' eye movements and their subjective impressions of the page's difficulty. Therefore, in research question R1, we believed that such a relationship could be modeled.

In R2, our visualization method is useful for a teacher when pages are difficult for students, according to Table 3. In addition, the relevance maps may help teachers revise teaching materials. Even if teachers use the gaze maps, the maps may not be able to help them because they are redundant. Teachers have to extract useful information from the gaze map. However, the relevance maps focused on some specific regions in our experiment. Note that our visualization method may not be useful for teachers when pages are not difficult. We can provide more useful information for teachers by removing easy pages in advance.

We believed the highlight maps accurately represented where students found difficulty. However, students did not always add highlights. The relevance maps could localize difficult contents even if students forgot to add highlights, because they were generated automatically from eye movements.

The major limitation of this study was that we did not consider explanations on each page, or the types of content (such as figures, text, and equations). We believe that such information can strongly affect students' reading behaviors and their subjective impressions of difficulty. To further investigate these details, analysis of the multimodal information will be needed, including images of pages and types of contents.

## 6. CONCLUSION

In this study, we proposed a method to model relationships between students' eye movements and their subjective impressions of the difficulty of pages. Our method generated relevance maps representing locations where students found difficulty, automatically based on their eye movements. Our experiment implied that the relevance maps could provide useful information for teachers when revising teaching materials, as the system suggests where to revise on each page based on the relevance maps.

We only used eye movement data for modeling the relationships. However, eye movement data alone cannot completely represent the contexts of teaching materials, such as the contents of each page. We believe that information could improve the quality of the relevance maps. For example, we may be able to use prior materials that were difficult for students to better understand equations, and much longer texts. In the future, we will combine eye movement data and additional information, such as the types of content on each page.

## ACKNOWLEDGEMENTS

We are grateful to Dr. Kaori Tamura for helping with data collection. This work was supported by JSPS KAKENHI Grant Number JP19K20421.

## REFERENCES

- De Koning, B. B. et al., 2010. Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and instruction*, Vol. 20, Issue 2, pp. 111-122.
- Jian, Y.-C., and Ko, H.-W., 2017. Influences of text difficulty and reading ability on learning illustrated science texts for children: An eye movement study. *Computers & Education* Vol. 113, pp. 263-279.
- Kingma, D. P., and Ba, J., 2015. Adam: A Method for Stochastic Optimization. *Proceedings of the 3rd International Conference for Learning Representations*. San Diego, US, pp. 7-9.
- Lapuschkin, S. et al., 2016. The LRP Toolbox for Artificial Neural Networks. *The Journal of Machine Learning Research*, Vol. 17, pp. 3938-3942.
- Mason, L. et al., 2013. An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education*, Vol. 81, Issue 3, pp. 356-384.
- Minematsu, T. et al., 2019. Analytics of Reading Patterns Based on Eye Movements in an e-Learning System. *Society for Information Technology & Teacher Education International Conference*, Las Vegas, NV, U.S., pp. 744-749.
- Nakamura, K. et al., 2008. Estimating learners' subjective impressions of the difficulty of course materials in e-learning environments. *APRU 9th Distance Learning and the Internet Conference*, Tokyo, Japan, pp. 199-206.
- Ogata, H. et al., 2015. E-book-based learning analytics in University education. *International Conference on Computer in Education (ICCE 2015)*, Hangzhou, China, pp. 401-406.
- Ohkawauchi, T. et al., 2012. Study of a Teacher Support Method For Improving Lectures for E-Learning on Demand. *Proceedings of the IEEE transactions on image electronics and visual computing workshop 2012*, Kuching, Malaysia, pp. 1-4.
- Shiino, T. et al., 2019. Page-wise Difficulty Level Estimation using e-Book Operation Logs. *Proceedings of the 9th International Learning Analytics and Knowledge Conference*. Tempe, U.S.
- Shimada, A. et al, 2015. Automatic Summarization of Lecture Slides for Enhanced Student Preview. *International Conference on Computer in Education (ICCE 2015)*, Hangzhou, China, pp. 218-227.
- The, B. and Mavrikis, M., 2016. A Study on Eye Fixation Patterns of Students in Higher Education Using an Online Learning System. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. Edinburgh, United Kingdom, pp. 408-416.